ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-494

Do the long-term impacts of North Carolina's early childhood programs depend on the quality of educational opportunity during school?

Robert C. Carr
Duke University

Tyler W. Watts
Teachers College, Columbia University

Jade M. Jenkins
University of California, Irvine

Yu Bai
Duke University

Ellen Peisner-Feinberg
The University of North Carolina at Chapel Hill

Helen F. Ladd
Duke University

Kenneth A. Dodge
Duke University

Prior research has found that financial investments in North Carolina's early childhood education programs—Smart Start and NC Pre-K—generated positive effects on student achievement in reading and mathematics through eighth grade (Bai et al., 2020). The current study examined if these effects were moderated by two dimensions of educational opportunity in NC public school districts, as measured by (1) the average academic achievement level in third grade and (2) the rate of growth in academic achievement from third to eighth grade. The Smart Start effect on eighth grade reading achievement was larger in school districts with higher levels of average achievement. Also, the NC Pre-K effect on eighth grade reading achievement was smaller in school districts with higher rates of achievement growth.

VERSION: December 2021

Suggested citation: Carr, Robert C., Tyler Watts, Jade M. Jenkins, Yu Bai, Ellen S. Peisner-Feinberg, Clara G. Muschkin, Helen F. Ladd, and Kenneth A. Dodge. (2021). Do the long-term impacts of North Carolina's early childhood programs depend on the quality of educational opportunity during school?. (EdWorkingPaper: 21-494). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/jrgc-8n94

Do the long-term impacts of North Carolina's early childhood programs depend on the quality of educational opportunity during school?

Robert C. Carr [a]

Tyler W. Watts [b]

Jade M. Jenkins [c]

Yu Bai [a]

Ellen S. Peisner-Feinberg [d]

Clara G. Muschkin [a]

Helen F. Ladd [a]

and

Kenneth A. Dodge [a]

[a] Duke University
[b] Teachers College, Columbia University
[c] University of California, Irvine
[d] The University of North Carolina at Chapel Hill

Author Note

Corresponding author: Robert C. Carr, Duke University, 2024 West Main Street, Duke Box 90539, Durham, NC 27708-0539. robert.carr@duke.edu.

Funding: This research study was supported by grant funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (1R01HD095930-01A1).

Declaration of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Acknowledgements: The authors wish to thank the North Carolina State Center for Health Statistics, the North Carolina Department of Public Instruction, the North Carolina Division of Child Development and Early Education, the North Carolina Partnership for Children, and the North Carolina Education Research Data Center for providing the relevant data files and guidance. Thank you to the members of the Consortium on Early Childhood Intervention Impact (CECII) for their helpful feedback on earlier versions of this work.

Abstract

Prior research has found that financial investments in North Carolina's early childhood education programs—Smart Start and NC Pre-K—generated positive effects on student achievement in reading and mathematics through eighth grade (Bai et al., 2020). The current study examined if these effects were moderated by two dimensions of educational opportunity in NC public school districts, as measured by (1) the average level of academic achievement among third-grade students in the school district and (2) the rate of growth in academic achievement among students in the school district as they progressed from third to eighth grade. The Smart Start effect on eighth grade reading achievement was larger in school districts with higher levels of average achievement. Also, the NC Pre-K effect on eighth grade reading achievement was smaller in school districts with higher rates of achievement growth.

Keywords: Child care; early childhood education; school quality; reading; mathematics; long-term effects.

Introduction

Public investments in children's early skill development can result in substantial benefits to society, especially if those investments target children from families who experience economic insecurity (Heckman, 2006). These benefits can include increased academic achievement and educational attainment, improved health outcomes, greater earning potential in adulthood, and reduced criminal behavior (Yoshikawa et al., 2013). Today, early childhood education (ECE) is widely recognized as a proven strategy to promote children's early skill development, with benefits that far outweigh their costs. For example, highly effective programs have generated returns on investments that range from three to seven dollars saved for every dollar invested (Yoshikawa et al., 2013).

Despite the obvious importance of investing in early childhood skill development, children also benefit from later investments in their skill development throughout childhood. Indeed, a wealth of evidence suggests that the quality of educational opportunity in schools can significantly promote children's skill development during childhood and adolescence (Eccles & Roeser, 2010). An important question not yet fully resolved is how early investments can interact with later investments in children's skill development. In particular, understanding if the quality of subsequent school environments moderates the long-term impacts of ECE throughout the formal schooling years is a topic of interest to both policy and scientific audiences alike (Phillips et al., 2017).[1]

The quality of the schooling environment has been hypothesized as a primary factor in determining whether the impacts of ECE persist in the long run (Bailey et al., 2017; Phillips et

[1] In the current study, we refer to ECE program effects on eighth grade outcomes as "long-term" effects because children have entered adolescence by eighth grade. Alternatively, use of the term "short-term" effects often pertains to effects on children's outcomes at the end of the preschool period and effects on child outcomes during elementary school are often referred to as medium-term effects (e.g., Unterman & Weiland, 2020).

al., 2017). Yet, there remains some debate regarding the mechanisms by which subsequent education quality will moderate the long-term impacts of ECE. There are currently two leading hypotheses: *dynamic complementarity* and *dynamic substitutability* (Bailey, Duncan, et al., 2020). Dynamic complementarity implies that ECE program impacts will be maintained and may even be enhanced in the context of higher-quality educational environments during school. Alternatively, dynamic substitutability implies that ECE program impacts will be diminished in the context of higher-quality educational environments during school. The existing research base is generally mixed and provides no clear consensus regarding the role of subsequent school environments in differentiating the impacts of ECE (Bailey, Jenkins, et al., 2020). Further research is needed to test these alternative hypotheses.

Previous research has found that financial investments in North Carolina's two, flagship ECE programs—Smart Start and NC Pre-K—produced positive effects on student academic achievement through eighth grade (described in greater detail below) (Bai et al., 2020).[2] In the current study, we examined whether these long-term effects were moderated by the quality of educational opportunity that children subsequently experienced in NC public school districts—focusing on program effects on student achievement in eighth grade. We used a linked administrative dataset including more than one million children who were born in NC across an 18-year period. Following Reardon's (2018) operationalization of "educational opportunity," we measured the quality of educational opportunity in NC public school districts along two dimensions: (1) the average level of academic achievement among third-grade students in the school district and (2) the rate of growth in academic achievement among students in the school

[2] The NC Pre-K program was previously known as the More at Four program. In 2011, the NC General Assembly transferred the existing More at Four program from the Department of Public Instruction to the Division of Child Development and Early Education in the Department of Health and Human Services and renamed it to the NC Pre-K program.

district as they progressed from third to eighth grade. We then tested for moderation of Smart Start and NC Pre-K program funding effects by the quality of educational opportunity in NC public school districts. The results of this study offer implications for the alignment of high-quality educational programming across the preschool and school-age periods.

North Carolina's ECE programs

Smart Start

The Smart Start program was established in 1993 by the NC General Assembly to create a comprehensive and integrated system of developmental services for children (and their families) from birth to age 4. The Smart Start program provides services in three main areas: child care (focusing on affordability, availability, and quality), family support, and child health. Smart Start is administered by the NC Partnership for Children, which awards funds to local Smart Start partnerships to provide program services in each of NC's 100 counties. Local decision-making is a hallmark of the Smart Start initiative as the activities of each local partnership are guided by a board of community members who select the services that best meet the needs of children and families in their community (North Carolina Partnership for Children, 2019). Therefore, the scope and availability of local Smart Start services can vary across counties and over time.

While some local Smart Start services are universally available to all children and families within communities, other services are targeted to children from lower-income families. For example, child care subsidies constitute a large portion of Smart Start expenditures and eligibility for the sliding fee scale is limited to families earning at or below 75% of the state's median income (North Carolina Partnership for Children, 2019). At the same time, efforts to improve the quality of child care programs in communities may benefit all children, regardless of

whether or not they receive a child care subsidy. During the time period of Smart Start implementation considered in our study, the quality of local child care centers sponsored by Smart Start was found to improve on a wide range of quality metrics (Bryant et al., 2003). For example, the percent of children attending high-quality child care programs (4- and 5-star programs rated on a 5-star scale) doubled from 33% in 2001 to 63% in 2010 (the last program year considered in our study) (North Carolina Partnership for Children, 2010).

NC Pre-K

Established by NC General Assembly in 2001, NC Pre-K is a state-funded educational program for eligible 4-year-old children, designed to enhance their school readiness skills prior to kindergarten entry. During our study period between 2001 and 2010, the NC Pre-K program was administered by the NC Department of Public Instruction (DPI) and then, in 2011, the NC General Assembly transferred the program to the Division of Child Development and Early Education (DCDEE) in the NC Department of Health and Human Services.[3] The state administrative agency awards funds to local contractors to provide NC Pre-K programming in each of NC's 100 counties—funding qualified child care centers to serve eligible children in a variety of classroom-based educational settings, including public schools, Head Start, and private child care centers (both for-profit and nonprofit). While NC was among the latter half of states to adopt a public pre-kindergarten program (Cohen-Vogel et al., 2020), NC Pre-K grew quickly to become an established program operating at scale. During the 2009-10 program year (the last program year considered in our study), 24% of the state's 4-year-old population participated in the program (N = 31,197; Barnett et al., 2010) and that figure has remained relatively consistent over time (e.g., 24% served in 2019; Friedman-Krauss et al., 2020).

[3] Previously known as the More at Four program, the program was renamed as the NC Pre-K program when it was transferred from the NC DPI to DCDEE.

NC Pre-K program guidelines allow 4-year-old children to receive NC Pre-K funding if their family has a gross family income at or below 75% of the state median income, or if the child has at least one of the following factors: educational or developmental delay, an identified disability, a chronic health condition, or limited English proficiency (Barnett et al., 2010).[4] Funding for NC Pre-K is targeted to children who meet the eligibility criteria. Nonetheless, aspects of the program's design may lead to benefits for other children in communities—particularly those children who do not directly receive NC Pre-K funding, but still participate in a NC Pre-K funded child care center through some other funding means. For example, NC Pre-K contractors are expected to meet a variety of program standards designed to ensure a high-quality, classroom-based educational experience for all children enrolled at their local sites—both NC Pre-K funding recipients and non-recipients. The state administrative agency regularly implements quality monitoring and improvement initiatives to promote uniformity in NC Pre-K program services across the state, to the extent possible. Therefore, all children in communities may benefit from increases in ECE quality because of NC Pre-K. Moreover, NC Pre-K has long maintained high standards for program quality. During the 2009-10 program year, the NC Pre-K program met all 10 of the quality benchmarks established by the National Institute of Early Education Research—representing an increase from 7 quality benchmarks met during the 2002-03 program year (Barnett et al., 2010; Barnett et al., 2003).

Background Literature

The Long-Term Impacts of ECE

A robust body of evidence suggests that ECE programming can have positive, short-term impacts on children's academic and social-behavioral readiness to succeed in elementary school

[4] After our study period, NC Pre-K program eligibility was extended to children with a parent serving in the military.

(Camilli et al., 2010). Research has also found that ECE programming can have long-term impacts on educational outcomes throughout the school-age years and into adulthood (McCoy et al., 2017). Early evidence comes from random assignment studies of two model ECE programs—the Abecedarian and Perry Preschool programs—which both documented positive effects on academic achievement outcomes at school entry and throughout the school-age years, as well as positive effects on overall educational attainment in adulthood; including higher rates of high school graduation and enrollment in postsecondary programs (Campbell et al., 2002; Schweinhart et al., 1985).

More recent studies have also documented evidence of the positive, long-term effects of modern day ECE programs operating at scale. This includes a series of quasi-experimental studies examining the long-term effects of financial investments in the rollout of the Smart Start and NC Pre-K programs (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2014; Muschkin et al., 2015). These effects include improvements to student achievement in reading and mathematics as well as reductions in special education placements and grade retention—effects that were evident in third grade and continued to persist into eighth grade. These findings were drawn from a research design that leveraged between-county and within-county variation in the rollout of funding for Smart Start and NC Pre-K across NC’s 100 counties over multiple years (a generalization of the difference-in-differences design). These effects were measured at a population-wide level and could have been conferred through direct effects on former program participants as well as spillover effects to non-participants.

Using similar research designs with county-level variation in ECE program rollout, several studies have documented evidence of the positive, long-term, and population-wide effects on educational outcomes based on the rollout of the federal Head Start preschool program

(Bailey et al., 2021; Johnson & Jackson, 2019; Kose, 2021; Thompson, 2018) and Georgia's universal pre-kindergarten program (Fitzpatrick, 2008), as well as the rollout of targeted preschool programming for low-income children in Denmark (Rossin-Slater & Wüst, 2020). Positive long-term effects of Head Start have also been documented based on studies of sibling comparisons (Currie & Thomas, 1995; Garces et al., 2002; Johnson & Jackson, 2019; Pages et al., 2020) and regression discontinuity research designs (Carneiro & Ginja, 2014; Ludwig & Miller, 2007). A majority of the aforementioned studies focused on program effects related to educational attainment outcomes in adulthood, such as high school graduation and overall years of schooling. However, two of these studies complement previous research on Smart Start and NC Pre-K program funding by documenting positive effects of ECE program funding on academic achievement outcomes during the school-age period (i.e., reading and math test scores; Fitzpatrick, 2008; Kose, 2021).

A separate body of research has used treatment-comparison group designs to contrast ECE and non-ECE program participants. Research in this area offers more extensive evaluation of ECE program effects on educational outcomes during the school-age period. Three random-assignment studies of modern day ECE programs operating at scale have documented positive short-term effects on students' school entry skills—including studies of the federal Head Start preschool program (Puma et al., 2010), Tennessee's state-funded pre-kindergarten program (Lipsey et al., 2018), and NC's NC Pre-K Program (Peisner-Feinberg et al., 2019). However, all of these studies document substantial fadeout in effects on academic achievement outcomes during the early elementary grades (Lipsey et al., 2018; Peisner-Feinberg et al., 2020; Puma et al., 2012). Surprisingly, the random-assignment study of Tennessee's state-funded pre-kindergarten program documented negative effects on academic achievement outcomes in the

early elementary grades—after finding initially positive effects at the end of preschool (Lipsey et al., 2018). Nonetheless, it is possible for ECE programming to demonstrate positive effects on longer-term educational outcomes, even if diminished or null effects on short-term academic achievement are observed (Bailey, Duncan, et al., 2020).

Studies based on quasi-experimental, treatment-comparison group contrasts also provide strong evidence of short-term ECE program effects on academic achievement outcomes at school entry (Barnett et al., 2018; Gormley et al., 2008; Peisner-Feinberg & Schaaf, 2011; Peisner-Feinberg et al., 2014; Weiland & Yoshikawa, 2013). Still, evidence of longer-term impacts on academic achievement outcomes is somewhat mixed in these studies—including evidence of fadeout in achievement effects documented during the early elementary grades (Bassok et al., 2019; Hill et al., 2015; Weiland et al., 2019) as well as persistent, positive effects on achievement outcomes during elementary school (Bassok et al., 2019; Hill et al., 2015), middle school (Gormley et al., 2018; Phillips et al., 2016), and high school (Barnett & Jung, 2021). Several of these quasi-experimental studies also provide evidence of ECE program effects on broader measures of educational attainment outcomes during the school-age period (Gormley et al., 2018; Gray-Lobe et al., 2021; Phillips et al., 2016), and even into college (Gray-Lobe et al., 2021). The effects of NC Pre-K have also been investigated in a series of quasi-experimental studies based on treatment-comparison group contrasts—documenting positive effects on academic achievement outcomes at school entry (Peisner-Feinberg & Schaaf, 2011) as well as positive effects on academic achievement and educational attainment outcomes during the early elementary grades, through third grade (Peisner-Feinberg et al., 2017; Peisner-Feinberg & Schaaf, 2010, 2011).

In sum, a robust body of evidence demonstrates that ECE programming can lead to improved academic achievement in the short term. Additionally, these effects can persist in the long term and may even lead to improvements in other educational attainment outcomes. This evidence is mainly drawn from random assignment studies of two model ECE programs implemented in the 1970’s as well as quasi-experimental studies of modern day ECE programs being implemented at scale. Several random assignment studies of modern day ECE programs document initially positive effects at school entry and subsequent fadeout in effects on academic achievement outcomes. However, it is not clear *if* and *how* the quality of subsequent school environments play a role in determining whether ECE program effects persist in the long run.

The Long-Term Impacts of ECE and Subsequent School Quality

There may be certain conditions under which ECE program impacts are more likely to persist in the long run. Moreover, these conditions may be complex and multifaceted, and their influence may vary across interventions and outcomes. One hypothesis suggests that ECE program impacts will be maintained or even enhanced in the context of higher-quality educational environments during school (Bailey et al., 2017; Cunha et al., 2006; Phillips et al., 2017). Commonly referred to as *dynamic complementarity*, this hypothesis is rooted in an economic theory of human capital investments and suggests that “early investments are not productive if they are not followed up by later investments” across the lifespan (Cunha et al., 2006, p. 698). In the education research literature, the *sustaining environments* hypothesis specifically identifies high-quality school environments as the subsequent investment that is needed to maintain the long-term impacts of ECE (Bailey et al., 2017). The fundamental premise of these complementary hypotheses rests on the idea that “skills beget skills” or “abilities beget abilities”—that skills acquired during an earlier stage of development persist into later stages of

development, and enhance the opportunity to acquire more advanced skills at later stages of development (Cunha et al., 2006, p. 703). Moreover, ECE program impacts on children's early skill development will carry over into long-term impacts on later skill development if higher-quality schooling enables children to further build on the early skills they gained through ECE and acquire more advanced skills during school. Through dynamic complementarity, the combination of ECE program exposure *and* high-quality schooling will lead to the highest skill attainment for children.

Alternatively, *dynamic substitutability* suggests that ECE program impacts will be diminished in the context of higher-quality educational environments during school, because the benefits of higher-quality schooling will substitute for the benefits of ECE program exposure (Bailey, Duncan, et al., 2020). Under these conditions, higher-quality schooling will lead to the highest skill attainment for all children regardless of ECE program exposure, as higher-quality school environments boost the skill development of children who missed out on ECE and enable them to catch up to their peers with previous ECE program exposure (Abenavoli, 2019). Dynamic substitutability implies that ECE program impacts will also be maintained in the context of *lower-quality* educational environments during school, as ECE program exposure supports children's continued skill development while compensating for *or* buffering children against subsequent lower-quality educational environments (Abenavoli, 2019; Bailey, Duncan, et al., 2020). For example, Abenavoli (2019) describes how high-quality ECE programming may promote resilience among children who subsequently enroll in lower-quality school environments—which is consistent with evidence that higher-quality child care can compensate for or buffer children against the negative impacts of lower socioeconomic status, including lower family income (McCartney et al., 2007). Importantly, dynamic substitutability only

implies that children with ECE program exposure will be better off relative to children without program exposure in lower-quality school environments (i.e., a positive impact of ECE), while all children will attain higher skills in higher-quality school environments regardless of ECE program exposure (Abenavoli, 2019; Bailey, Duncan, et al., 2020). Statistically, support for the dynamic substitutability hypothesis would be evidenced by a negative coefficient on the interaction between ECE program exposure and subsequent schooling quality; suggesting a counterpoint to the dynamic complementarity hypothesis which would be evidenced by a positive interaction coefficient between ECE and subsequent schooling (Bailey, Jenkins, et al., 2020).

There are generally mixed findings across studies that have examined how subsequent school contexts moderate the long-term effects of ECE (see review by Bailey, Jenkins, et al., 2020). With regard to financial investments in ECE programming, Johnson and Jackson (2019) found evidence in support of dynamic complementarity, whereby the long-term benefits of Head Start investments were larger in the context of subsequent higher investments in public schooling. Other studies based on ECE treatment-control group designs also provide support for dynamic complementarity. These studies considered a variety of subsequent education quality measures, including teacher quality (Swain et al., 2015), school-wide academic achievement (Ansari & Pianta, 2018; Carr et al., 2021; Unterman & Weiland, 2020; Zhai et al., 2012), and school-wide growth in academic achievement (Carr et al., 2021; Pearman et al., 2020; Unterman & Weiland, 2020). One study in particular examined the effects of NC Pre-K program participation on child outcomes at the end of kindergarten and found that program effects on child language and working memory skills were only evident in higher-quality elementary schools as measured by school-wide academic proficiency and academic growth, respectively

(Carr et al., 2021). However, similar effects on child literacy and math outcomes were not evident in that study.

Alternatively, several studies of ECE program participation provide support for the dynamic substitutability hypothesis of diminished ECE impacts in higher-quality school contexts (Bierman et al., 2014; Curenton et al., 2015; Magnuson et al., 2007). Perhaps most prominently, Magnuson et al. (2007) found that positive effects of ECE participation on academic achievement in elementary school were diminished in the context of smaller class sizes and more time spent in reading instruction, but ECE effects were maintained during elementary school for students in larger classes with less reading instruction (i.e., lower-quality environments).

The Quality of Educational Opportunity in Public Schools

In the current study, we sought to index the quality of educational opportunity in NC’s public schools. We follow Reardon (2018) who defined and measured two dimensions of educational opportunity by using standardized test score data from all public-school districts in the U.S. The first dimension was measured by the average achievement level among third grade students in the school district (i.e., average achievement), while the second dimension was measured by the rate of growth in academic achievement among students in the school district as they progressed from third to eighth grade (i.e., achievement growth). Reardon (2018) found that average achievement was highly correlated with a measure of community-wide socioeconomic status (r = 0.68), while achievement growth was only moderately correlated (r = 0.32). These correlations suggest that the two measures capture distinct dimensions of educational opportunity. While average achievement may reflect a broad range of *community-wide educational opportunities* available to students (e.g., neighborhood characteristics, average family resources in a community), achievement growth may better capture the unique

contributions of schooling toward promoting student achievement during the school-age years. Our study utilized Reardon's (2018) framework to examine these dimensions of educational opportunity among children attending NC public schools—a school system where student achievement measures are commonly used for high-stakes accountability and quality monitoring. Following Reardon (2018), we also examined school districts rather than schools as the unit of analyses for educational opportunity. There are many ways in which school districts can directly and indirectly shape the quality of educational opportunity that students experience in their schools.

Case studies demonstrate that school-district administrators make important decisions that directly influence the structural characteristics and culture of schools, including decisions about curriculum; decisions about hiring, staffing, and salaries of teachers and principals; as well as decisions about the professional development and evaluation of teachers and principals (Chenoweth, 2017, 2021). In extreme cases, school district administrators can fundamentally reform and/or reorganize schools through school closures, consolidations, and the implementation of accountability systems. School-district administrators can also influence the culture of school environments. For example, Chenoweth (2021) described effective school districts as having a culture in which the leadership cultivates the belief that all children are capable of success and that the adults are responsible for ensuring children succeed (p. 131). Indeed, a comprehensive case study of Chicago Public Schools identified *effective leaders* as the most critical ingredient in building a culture of school improvement (Bryk et al., 2010). In sum, school-district leadership can shape the structural organization and culture of schools, and, in turn, these features of the school environment can influence the quality of educational opportunity children experience (Eccles & Roeser, 2010).

The Current Study

The quality of educational environments during school may serve to enhance or diminish the impacts of ECE as children progress through school. However, the extant evidence is mixed regarding *if* and *how* the quality of subsequent school environments will play a role in the persistence of ECE program impacts in the long run (Bailey, Jenkins, et al., 2020). The current study sought to address this issue by examining whether the long-term effects of North Carolina's Smart Start and NC Pre-K programs were moderated by the quality of educational opportunity in NC public schools. In the first phase of our study, we used a growth curve model approach to index the quality of educational opportunity among students in NC public school districts by (1) the average level of academic achievement among third-grade students in the school district and (2) the rate of growth in academic achievement among students in the school district as they progressed from third to eighth grade.[5] In the second phase of our study, we considered variation in Smart Start and NC Pre-K program funding allocated to NC's 100 counties across 18 program years to examine the effects of program investments on student reading and mathematics outcomes in eighth grade—consistent with the analysis framework used in previous studies (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2014; Muschkin et al., 2015).[6] We simultaneously tested for moderation of Smart Start and NC Pre-K program funding effects by the quality of educational opportunity in NC public schools. In particular, we tested alternative hypotheses in relation to the moderation of Smart Start and NC Pre-K program funding effects. Support for the dynamic complementarity hypothesis would be evidenced by

[5] We considered average academic achievement level in third grade because it was the lowest grade level at which standardized achievement tests were administered in NC.

[6] The current study extended the number of years of Smart Start and NC Pre-K program funding considered in the previous studies by six additional years to include 18 fiscal years of funding information. In doing so, we also increased the sample size by 303,803 students for the reading outcome analyses and 265,390 students for the math outcome analyses.

positive effects of program funding and educational opportunity as well as a positive interaction between these variables. Support for the dynamic substitutability hypothesis would be evidenced by positive effects of program funding and educational opportunity as well as a negative interaction between these variables.

Methods

Data Sources

Our study used data on the full population of students enrolled in NC public schools in third grade through eighth grade between the 1996 and 2019 school years[7] (school records obtained from the NC DPI) as well as a subpopulation of these students who were also born in NC between January 1, 1987, and August 31, 2005 (birth records obtained from NC State Center for Health Statistics). Data on the full population of students were used in the Phase I analyses while data on the subpopulation of students were used in the Phase II analyses (described in further detail in the *Analyses* section below). To combine these data, the identified birth records and school records of individual students were matched by the NC Education Research Data Center (NCERDC) at Duke University. In the current study, 74% of all birth records were matched to a school record. We also utilized on data on state allocations for Smart Start program funding obtained from the NC Partnership for Children, data on state allocations for NC Pre-K program funding obtained from the NC DPI, and publicly available data on each of NC's 100 counties obtained from the NC Office of State Budget and Management (OSBM). Institutional Review Board approval for this research was obtained at Duke University under the project titled "Combining Birth Data with Longitudinal Data on Schooling to Explore the Relationships

[7] A school year is defined as July 1st of the previous year through June 30th of the listed year.

Between Children's Birth Weight, Immigrant Status, Pre-School Experiences and Performance in School” (Protocol: 2017-0495).

Measures

Smart Start funding exposure

Data on Smart Start funding were obtained from the NC Partnership for Children. Smart Start was established during state fiscal year 1994 and funds were allocated to program contractors on a state fiscal year basis.[8] We converted all funds into 2019-dollar values based on the Consumer Price Index. To calculate exposure to Smart Start, we divided the amount of Smart Start funds awarded to each county during each fiscal year by the number of age-eligible children (i.e., ages 0–4) in each county during each fiscal year (based on county-level population estimates from NC OSBM).[9] To calculate cumulative exposure to Smart Start, we summed the annual Smart Start exposure in each county across each five-year period in which children in our panel were eligible to benefit from the program (i.e., prior to kindergarten entry between ages 0–4). After the program was established in 1994, the average county-level, cumulative exposure to Smart Start funding grew steadily and reached a peak of $2,383 in 2004, but then declined to $1,853 in 2010 (the last year of program funding considered in the current study) (see Figure 1). The average county-level, cumulative exposure to Smart Start funding between program years 1994-2010 was $1,539.

[8] The NC Partnership for Children provided Smart Start funding data expended by program contractors in each of NC’s 100 counties during each fiscal year.
[9] Note the Smart Start and NC Pre-K funding exposure variables were based on the entire population of age-eligible children even if only a portion of children received direct financial support to participate in the programs.

NC Pre-K funding exposure

Data on NC Pre-K funding were obtained from the NC DPI. NC Pre-K was established during state fiscal year 2002 and funds were allocated to program contractors on a state fiscal year basis.[10] We converted all funds into 2019-dollar values based on the Consumer Price Index. To calculate exposure to NC Pre-K, we divided the amount of NC Pre-K funds awarded to each county during each fiscal year by the number of age-eligible children (i.e., age 4) in each county during each fiscal year (based on county-level population estimates from NC OSBM). After the program was established in 2002, the average county-level exposure to NC Pre-K program funding grew steadily and reached a peak of $2,030 in 2010, with high variance within and across counties over the study period (see Figure 2). The average county-level exposure to NC Pre-K program funding between program years 2002-2010 was $1,064. We estimated that the program directly served 28% of all four-year-old children in NC in 2010 (the last year of program funding considered in the current study).

Student Achievement in Reading and Mathematics

Standardized assessments of reading and mathematics achievement were administered to students in NC public schools by school administrators at the end of each grade between third grade and eighth grade. Our study used scores from these end-of-grade (EOG) tests administered during all school years between 1996 and 2019.[11, 12] These data were obtained from the NCERDC. The EOG tests are state-mandated, standardized assessments with well-documented

[10] A total of 91 NC Pre-K program contractors served NC's 100 counties during each fiscal year. Funds awarded to a single contractor that served multiple counties were disaggregated across counties within the contractor's service region based on the population of age-eligible children in each county.
[11] The test scores of students who took alternative assessments designed for students with disabilities were excluded from these analyses because those test scales were not equated with the EOG scale.
[12] A number of advanced eighth grade students were not assessed with the EOG Math test in 2019, but were assessed using the high-school end-of-course (EOC) math test instead (1.6% of our input sample). These students were excluded from our analyses.

psychometric characteristics that satisfy the requirements of the U.S. Department of Education (Bazemore & Van Dyk, 2004; Mbella et al., 2016a, 2016b; North Carolina Department of Public Instruction, 2009; Sanford, 1996).

Separate editions of the EOG tests were administered across the range of years included in our panel study (i.e., four editions of the EOG Reading and five editions of the EOG Math). To account for changes in the scaling of EOG test editions across years, we converted the EOG scores into *Z*-scores (across all editions) by subject, school year, and grade, resulting in a *Z*-score of 0.00 equal to the statewide average for students in each grade during each school year.

Covariates

Student characteristics. Data on individual students were obtained from (a) the birth records data and (b) the school records data. The student covariates include student sex (Female = 0, Male = 1; school records); student birth weight categorized as extremely low, very low, low, normal (reference category), or high (birth records); student race categorized as African American, Native American, Asian, Hispanic, mixed race, or non-Hispanic White (reference category) (school records); maternal education in years (birth records); parent marital status (Not married = 0, Married = 1; birth records); maternal age in years (birth records); no dad information on birth record (birth records); mother immigrant (No = 0, Yes = 1; birth records); student was first born (No = 0, Yes = 1; birth records); mother race categorized as African American, Native American, Asian, Hispanic, other race, or non-Hispanic White (reference category) (birth records); and the quarter of the school year in which the child was born categorized as quarter 1, 2, 3 (reference category; defined as the youngest children eligible for kindergarten entry during the school year), and 4.

School-district characteristics. Data on NC's public-school districts across years were obtained from the NCERDC. The school-district covariates include the per pupil expenditures from local, state, and federal funding sources (converted to 2019 $s), the percent of economically disadvantaged students in the school district, and the total number of students in the school district's membership.[13]

County characteristics. Data on each of NC's 100 counties across years were obtained from (a) the NC OSBM data or (b) the birth records data aggregated from student-level information to county-level information. The county covariates include the percent of births to black mothers, Hispanic mothers, and low education mothers (< 12 years of education) (birth records); the percent of population receiving Food Stamps and the percent of population receiving Medicaid (NC OSBM); the total number of births (log transformed; birth records); the total population (log transformed; NC OSBM); and median family income (converted to 2019 $s; NC OSBM).

Analyses

Analyses proceeded in two phases. In the first phase, we indexed the quality of educational opportunity in NC Public Schools by measuring school-district average achievement and achievement growth among students from third to eighth grade. Based on the resulting measures of average achievement and achievement growth, we examined associations of these measures. In the second phase of analysis, we examined the long-term effects of Smart Start and NC Pre-K program funding exposure on student achievement outcomes in eighth grade, and we tested whether these long-term effects were moderated by school-district average achievement and achievement growth. All analyses were completed in SAS® version 9.4.

---

[13] Economic disadvantage was indexed based on the percent of students in the school district who qualified for free- or reduced-price lunch.

Phase I: Measuring School-District Average Achievement and Achievement Growth

In Phase I of our analyses, we measured school-district average achievement and achievement growth for 18 cohorts of students who progressed from third to eighth grade in each of NC's 115 public school districts (see Figure S1).[14, 15,16] We used the reading and math scores drawn from the full population of students enrolled in NC public schools. The growth cohorts were defined as all students who took the third grade through eighth grade reading and/or math assessments during a six-year period.[17, 18] For example, the first growth cohort was defined by all students who took the third-grade reading and/or math assessment in 1996, the fourth-grade assessment in 1997, the fifth-grade assessment in 1998, the sixth-grade assessment in 1999, the seventh-grade assessment in 2000, and/or the eighth-grade assessment in 2001.

We estimated a multi-level growth curve model for each of the 18 growth cohorts, with separate models estimated for reading and math scores ($18 \times 2 = 36$ models total). All models were estimated using the HPMIXED procedure in SAS® version 9.4. The generalized equation for the multi-level growth curve models to measure school-district average achievement and achievement growth is shown below (with separate models calculated for each of the 19 cohorts):

[14] We note that 4 public school districts were closed and consolidated with other public school districts during the course of our panel years. For our study, we aggregated information within these consolidated school districts prior to the years in which they were consolidated to utilize 115 school districts across all study years.
[15] All charter school districts were excluded from our study. The following public-school districts were also excluded from our study because they were not comparable to the 115 traditional public-school districts: (1) the NC Schools for the Blind and Deaf and (2) the NC School of Science and Mathematics.
[16] There are 75 counties in NC that were served by a single school district, while 15 counties were served by two school districts.
[17] The scores of individual students who were retained in grade and had valid test scores recorded for the same grade assessment in separate years were included in the models for multiple cohorts (e.g., the scores for a student who took the third-grade assessment during the 1997 and 1998 school years were included in the models for cohorts 2 and 3, respectively).
[18] Within cohorts, the scores of students who changed school districts were included in each respective school district (e.g., the scores for a student who took the third- through fifth-grade assessments in school district A were nested within school district A, and the scores for that same student who took the six- through eighth-grade assessments in school district B were nested within school district B).

Level-1 (time; *t*):

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}time_{tij} + e_{tij}$$

Level-2 (students; *i*):

$$\pi_{0ij} = \beta_{00j} + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + r_{1ij}$$

Level-3 (school districts; *j*):

$$\beta_{00j} = \gamma_{000} + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + u_{10j}$$

Variance Components:

$$e_{tij} \sim N(0, \sigma_{tij})$$
$$r_{0ij} \sim N(0, \sigma_{0ij})$$
$$r_{1ij} \sim N(0, \sigma_{1ij})$$
$$u_{00j} \sim N(0, \sigma_{00j})$$
$$u_{10j} \sim N(0, \sigma_{10j})$$

$\boldsymbol{Y_{tij}}$ is the test score at time *t* (grade) for student *i* in school district *j*; $time_{tij}$ is coded as 0 at third grade, 1 at fourth grade, 2 at fifth grade, 3 at sixth grade, 4 at seventh grade, and 5 at eighth grade (separate models were calculated for reading and mathematics scores)

$\boldsymbol{\pi_{0ij}}$ is the expected achievement level (i.e., intercept) for student *i* in school district *j* in third grade

$\boldsymbol{\pi_{1ij}}$ is the expected achievement growth (i.e., linear slope) for student *i* in school district *j* during each grade between third and eighth grade

$\boldsymbol{\beta_{00j}}$ is the expected achievement level (i.e., intercept) for students in school district *j* in third grade

$\boldsymbol{\beta_{10j}}$ is the expected achievement growth (i.e., linear slope) for students in school district *j* during each grade between third and eighth grade

In this model, repeated assessments of students' reading or mathematics skills over time (*t*; Level-1) were nested within students (*i*; Level-2), and students were nested within school districts (*j*; Level-3). Time was coded as 0 at third grade, 1 at fourth grade, 2 at fifth grade, 3 at sixth grade, 4 at seventh grade, and 5 at eighth grade. Test scores (Y_{tij}) were modeled as a function of (1) an intercept term centered at the third-grade assessment score to represent the expected achievement level for student *i* in school district *j* in third grade (π_{0ij}) and the expected achievement level for students in school district *j* in third grade (β_{00j}) as well as (2) a linear

slope term to represent the expected achievement growth for student *i* in school district *j* during each grade between third and eighth grade (π_{1ij}) and the expected achievement growth for students in school district *j* during each grade between third and eighth grade (β_{10j}). The intercept term was allowed to vary randomly between students (r_{0ij}) and school districts (u_{00j}), and the slope term was also allowed to vary randomly between students (r_{1ij}) and school districts (u_{10j}). All of the variance terms were parameterized to be normally distributed random variables with a mean of zero. We specified an unstructured covariance matrix to allow the random intercepts and slopes and their variances to be correlated with one another at Level-2 and at Level-3 in order to allow for a systematic relation between intercepts and slopes. We used restricted maximum likelihood (REML) to estimate the variance components and the "residual" method to calculate denominator degrees of freedom. Students with at least one test score were included in these analyses and maximum likelihood was used to account for missing data (Singer & Willett, 2003). No covariates were included in these analyses in order to calculate unconditional estimates of the intercepts and slopes. Based on the results of this model, we derived estimates of average achievement (i.e., intercept; β_{00j}) and achievement growth (i.e., slope; β_{10j}) in reading/math for each of the 115 school-districts across the 18 cohorts for use in the subsequent analyses. Specifically, we derived the empirical best linear unbiased predictions (EBLUPs) for the realizations of the random intercept, slope, and nested errors. This model was conceptually similar to model 8.15–8.17 described by Bryk and Raudenbush (1992). Based on results of these analyses, we examined associations in order to better understand the validity of these measures, including associations between school-district economic disadvantage and average achievement/achievement growth.

Phase II: Eighth-Grade Student Achievement Analyses

In Phase II of our analyses, we examined the effects of Smart Start and NC Pre-K program funding exposure on student reading and math achievement outcomes in eighth grade as well as moderation of these effects by school-district average achievement and achievement growth. Using two-way fixed effect regression analyses, we estimated program funding effects as the weighted average of between-county effects and within-county effects across time—weighted by variance and sample size (Kropko & Kubinec, 2020). A generalized equation is shown below for the two-way fixed effect analyses to examine moderation of Smart Start and NC Pre-K Program effects on students' eighth grade outcomes by school-district average achievement and achievement growth:

$$\begin{gathered} O_{iscByPySy} = \beta_0 + \beta_1 SS\ funds_{cPy} + \beta_2 NCPK\ funds_{cPy} \\ +\beta_3 AvgAch_{scSy} + \beta_4 SS\ funds\ x\ AvgAch_{scSy} + \beta_5 NCPK\ funds\ x\ AvgAch_{scSy} \\ +\beta_6 AchGrowth_{scSy} + \beta_7 SS\ funds\ x\ AchGrowth_{scSy} + \beta_8 NCPK\ funds\ x\ AchGrowth_{scSy} \\ + \boldsymbol{X}_i + \boldsymbol{Z}_{scSy} + \boldsymbol{C}_{cBy} + \alpha_c + \gamma_{Py} + \kappa_{Sy} + \alpha * \kappa_{cSy} + \epsilon_{iscByPySy} \end{gathered}$$

In this model, O is the eighth-grade test score (with separate models estimated for reading and mathematics scores) for student *i*, in school district *s*, in county *c*, born in birth year *By*, exposed to program funding during program year *Py*, and enrolled in eighth grade during school year *Sy*; β_1 is the main effect of Smart Start funding exposure; β_2 is the main effect of NC Pre-K funding exposure; β_3 is the main effect of school-district average achievement (i.e., average achievement in reading for the reading outcome model and average achievement in math for the math outcome model); β_4 is the interaction between Smart Start funding exposure and school-district average achievement; β_5 is the interaction between NC-Pre-K funding exposure and school-district average achievement; β_6 is the main effect of school-district achievement growth (i.e., achievement growth in reading for the reading outcome model and achievement growth in

math for the math outcome model); β_7 is the interaction between Smart Start funding exposure and school-district achievement growth; β_8 is the interaction between NC-Pre-K funding exposure and school-district achievement growth; X is a vector of effects for the student-level covariates; Z is a vector of effects for the school-district-level covariates; C is a vector of effects for the county-level covariates; α is the fixed effect for county of residence to model between-county variation in program funding; γ is the fixed effect for program year to model within-county variation in program funding over time; κ is the fixed effect for school year to model variability in eighth-grade student achievement outcomes over time; and $\alpha * \kappa$ is the linear time trend for counties across school years to account for possible preexisting time trends in student eighth-grade achievement outcomes within counties over time (for more details, see Vincent, 2018) (see Figures S2 & S3 for county-level means of eighth grade reading and math *Z*-scores). We also examined if our findings were robust to the inclusion of quadratic time trends ($\alpha * \kappa * \kappa_{cSy}$). The program funding, county-level covariates, and county fixed effect were aligned to students' county of residence at birth.[19, 20] The school-district average achievement and achievement growth measures, and the school district covariates were aligned to the students' school district and school year during eighth grade. The models were estimated with robust standard errors clustered at the county level. The dependent variables and continuous independent variables were standardized for the analysis sample with *M* = 0, *SD* = 1 (with separate standardization conducted for the reading and math outcome analysis samples).

[19] In previous studies, we have found the pattern of findings remained the same between analyses based on county of residence at birth and county of school attendance (Bai et al., 2020).
[20] We conducted a series of endogeneity analyses to examine the association between Smart Start & NC Pre-K program funding exposure and school-district average achievement & achievement growth (see Table S1). We found no reliable association between Smart Start & NC Pre-K funding and school district average achievement or achievement growth in reading and math. These findings suggest that our analyses satisfy a necessary condition to examine moderation.

Dichotomous independent variables were centered, but not standardized. All models were estimated using the SURVEYREG procedure in SAS® version 9.4.

Based on the subpopulation of students with matched birth records and school records, we applied listwise deletion of cases with missing data on one or more study variables. From the matched sample, only 3% and 6% of students had missing data for the reading and math outcome analyses, respectively. For students who repeated eighth grade, we utilized only the first test score that was recorded during the first school year in which they were enrolled in that grade. Additionally, we implemented a “current year minus 1” approach to assign school-district average achievement and achievement growth values to students in a given school year on the basis of values derived from the prior school year, with a different set of students (e.g., students enrolled in eighth grade in 2002 were assigned school-district average achievement and achievement growth values from cohort #1, which was comprised of students enrolled in eighth grade in 2001, and so on).[21] This approach was similar to that used by Andrabi et al. (2011).

Results

Table 1 summarizes the descriptive information on all study variables at the student level.

Phase I: Measuring School-District Average Achievement and Achievement Growth

We measured school-district average achievement and achievement growth in reading and mathematics for 18 cohorts of students who progressed from third to eighth grade between 1996 and 2018. Based on EOG scores transformed to *Z*-scores, we found that school-district average achievement in third grade was –0.13 (*SD* = 0.22) for reading and –0.14 (*SD* = 0.23) for

[21] This approach was employed in order to avoid the confound using student test score information from the same students as both independent and dependent variables, because regressing student’s eighth grade achievement scores on school-district average achievement and achievement growth values measured from the same student test scores could introduce the same estimation errors on both the left- and right-hand side of the regression equation—resulting in biased estimates for school-district average achievement and achievement growth.

math (see Table S2). These findings indicated that the average third grade achievement of students in the public-school districts in our sample was slightly more than one tenth of a standard deviation below the statewide average.[22] Based on EOG scores transformed to *Z*-scores, we found that school-district achievement growth from third to eighth grade was –0.02 points (SD = 0.20) for reading and –0.01 points (SD = 0.20) for math.[23] These findings indicated that, on average, students in the public-school districts in our sample showed only slight declines in their statewide rank-order standing in *Z*-scores between third and eighth grade.

We then examined associations between school-district average achievement/achievement growth and school-district economic disadvantage. We found that school-district average achievement in reading and math was negatively associated with school-district economic disadvantage (β = –0.46, $p < .001$; β = –0.39, $p < .001$) (see Table S3). Alternatively, we found that school-district achievement growth in reading and math was not reliably associated with school-district economic disadvantage (β = –0.09, $p = .21$; β = –0.09, $p = .26$). Consistent with findings by Reardon (2018), these findings suggest that average achievement may reflect a broad range of socioeconomic conditions of the community where the school district is located, while achievement growth may better reflect the unique contributions

[22] Please note that students in charter school districts were included in the transformation of EOG scale scores to *Z*-scores, but these students were excluded from Phase I and II analyses. Therefore, a *Z*-score value of 0.00 indicates the statewide average for students in public school districts and charter school districts. The negative mean values for school-district average achievement and achievement growth imply that students in charter school districts scored higher than students in public school districts.

[23] Values for school-district achievement growth were rescaled to represent the rate of growth from third to eighth grade by multiplying the yearly rate of growth produced in the Phase I analyses by a value of 5. The interpretation for a school-district achievement growth score of 0.00 is that the school district's rank-order standing remained at the statewide average from third to eighth grade. The interpretation for a positive school-district achievement growth score is that the school district gained in rank-order standing among all public-school districts from third to eighth grade, while the interpretation for a negative school-district achievement growth score is that the school district declined in rank-order standing among all public-school districts from third to eighth grade.

of schooling toward promoting student achievement during the elementary and middle school grades.

Phase II: Eighth-Grade Student Achievement Analyses

Table 2 summarizes the results of regression analyses to examine the effects of Smart Start and NC Pre-K funding exposure on students' eighth-grade reading and math achievement as well as moderation of these program effects by each measure of school-district educational opportunity. Results are presented separately for student reading and math achievement outcomes. Within each outcome measure, separate results are presented for analyses with linear and quadratic time trends.

Reading. We found positive effects of school-district average achievement in reading ($\beta = 0.062$, $p < .001$; Model 1) and achievement growth in reading ($\beta = 0.037$, $p < .001$) on student's eighth grade reading achievement. While the main effect of Smart Start funding exposure was not statistically significant ($\beta = 0.012$, $p = .097$), the effect was moderated by school-district average achievement such that the magnitude of the Smart Start effect increased as school-district average achievement increased ($\beta = 0.007$, $p = .002$). Moreover, when we probed the interaction, we found that the Smart Start effect was statistically significant at *high* levels of school-district average achievement (i.e., at +1 *SD*, $\beta = 0.019$, $p = 0.01$), but was not statistically significant at *average* (i.e., at the mean, $\beta = 0.012$, $p = 0.10$) and *low* levels of school-district average achievement (i.e., at −1 *SD*, $\beta = 0.005$, $p = 0.53$) (see Figure 3). We also found a positive main effect of NC Pre-K funding exposure ($\beta = 0.017$, $p = .03$) and the effect was moderated by school-district achievement growth such that the magnitude of the NC Pre-K effect decreased as school-district achievement growth increased ($\beta = –0.003$, $p = .04$). Moreover, when we probed the interaction, we found that the NC Pre-K effect was not

statistically significant at *high* levels of school-district achievement growth (β = 0.014, p = 0.08), but was statistically significant at *average* (β = 0.017, p = 0.03) and *low* levels of school-district achievement growth (β = 0.020, p = 0.01) (see Figure 4). These findings were robust to the inclusion of quadratic time trends.

Math. We found positive effects of school-district average achievement in math (β = 0.083, p < .001) and achievement growth in math (β = 0.093, p < .001) on student's eighth grade math achievement and these findings were robust to the inclusion of quadratic time trends. The main effects of Smart Start funding exposure (β = 0.009, p = .29) and NC Pre-K funding exposure (β = 0.020, p = .09) were not statistically significant and we found no evidence of moderation based on analyses with linear time trends. However, additional findings emerged with the inclusion of quadratic time trends. These findings include a positive main effect of Smart Start funding exposure (β = 0.035, p = .002) as well as moderation of the Smart Start effect by school-district average achievement such that the magnitude of the Smart Start effect increased as school-district average achievement increased (β = 0.006, p = .02) and moderation of the Smart Start effect by school-district achievement growth such that the magnitude of the Smart Start effect increased as school-district achievement growth increased (β = 0.007, p = .01). Additionally, the NC Pre-K effect was moderated by school-district average achievement such that the magnitude of the NC Pre-K effect decreased as school-district average achievement increased (β = –0.005, p = .04).

Discussion

High-quality ECE can have long-term impacts on student educational outcomes throughout school and into adulthood (McCoy et al., 2017). However, what remains unclear is *if* and *how* these long-term impacts will be moderated by the quality of subsequent educational

environments during school (Bailey, Jenkins, et al., 2020). In the current study, we examined the eighth-grade effects of financial investments in NC's two, flagship ECE programs—Smart Start and NC Pre-K—and we simultaneously examined moderation of these effects by the quality of educational opportunity in NC public school districts. We examined two hypotheses regarding the mechanisms by which subsequent education quality may moderate the long-term effects of ECE—*dynamic complementarity* and *dynamic substitutability*. Interestingly, we found support for dynamic complementarity in relation to the Smart Start program effect and we found support for dynamic substitutability in relation to the NC Pre-K program effect.

With regard to the Smart Start effect on students' eighth grade reading achievement, we found that increases in program funding exposure had a positive effect, but this effect was only evident in the context of higher-quality school districts—as measured by the average level of reading achievement among third-grade students in the school-district. This finding provides support for the dynamic complementarity hypothesis and suggests that Smart Start program effects are likely to be larger in the context of higher-quality schooling environments.

With regard to the NC Pre-K effect on students' eighth grade reading achievement, we found that increases in program funding exposure had a positive effect, and this effect was only evident in the context of average-quality and lower-quality school districts, but not higher-quality school districts—as measured by the rate of growth in academic achievement among students in the school district as they progressed from third to eighth grade. This finding provides support for the dynamic substitutability hypothesis and suggests that NC Pre-K program effects are likely to be smaller in the context of higher-quality schooling environments. It is important to note that this finding revealed that all students in high-growth school districts were more likely to attain higher reading achievement in eighth grade compared to students in lower-growth

school districts, regardless of their level of NC Pre-K program funding exposure. Therefore, the developmental benefit of higher-growth school districts appeared to act as a substitute for children with low *or* no exposure to NC Pre-K during early childhood. Simultaneously, higher exposure to NC Pre-K appeared to buffer children against the ill effect of subsequently enrolling in a lower-growth school district—the reading achievement of students in lower-growth school districts more closely approximated those of students in higher-growth school districts if they were also exposed to higher-levels of NC Pre-K program funding.

Regarding the Smart Start and NC Pre-K effects on students' eighth grade math achievement, we found evidence of program effects and moderation that was not robust to different analytic specifications (i.e., modeling linear or quadratic time trends). However, these additional findings provided further evidence of dynamic complementarity between Smart Start program exposure and subsequent educational opportunity as well as dynamic substitutability between NC Pre-K program exposure and subsequent educational opportunity.

We speculate about the reasons why different dynamic effects emerged for Smart Start and NC Pre-K. First, scholars have suggested that dynamic substitutability is more likely to occur when ECE and subsequent schooling "are redundant in their content" (Bailey, Duncan, et al., 2020, p. 68). Dynamic substitutability could be evident for NC Pre-K because the program has long maintained a focus on direct instruction related to school readiness skills, including frequent and high-quality instruction related to early literacy building (e.g., Peisner-Feinberg et al., 2013). Moreover, observations of pre-kindergarten classrooms across 11 states (including data from NC) has also confirmed that language/literacy instruction is the most frequent learning activity observed in pre-kindergarten classrooms, constituting an estimated 17% of the entire day (Early et al., 2010). Hence, in communities with low or no exposure to NC Pre-K, enrolling in a

school district with higher rates of growth in reading achievement may have enabled students to make up the gains in underlying literacy skills that NC Pre-K would have otherwise conferred.

Alternatively, dynamic complementarity could be evident for Smart Start because the program's goals were related to broad improvements in the fundamental infrastructure of ECE—including child care affordability, availability, and quality, as well as family support and child health—rather than an explicit focus on direct instruction related to early literacy skills, akin to the focus provided by NC Pre-K. Nonetheless, Smart Start's broad mission may have served to promote student literacy achievement in the long run through means other than direct instruction in early literacy skills, such as enhanced family engagement in book reading. Therefore, in communities with high exposure to Smart Start, enrolling in a school district with higher levels of reading achievement may have enabled students to further build on the broad developmental benefits conferred through Smart Start. In sum, it is possible that we observed dynamic substitutability for NC Pre-K because the benefits conferred through its specific intervention components (e.g., early literacy instruction) closely corresponded to the benefits conferred through literacy instruction in high-growth school districts. In contrast, the specific intervention components characteristic of Smart Start programming might confer benefits that lay a broader foundation for high-quality schooling to build on, so we see dynamic complementarity.

In the debate about fadeout or persistence of ECE program effects on student achievement, the findings of our research should encourage educational stakeholders and policymakers to consider how subsequent schooling environments play a role in determining whether program effects are maintained or diminished in the long run. While previous studies have documented the enduring effects of Smart Start and NC Pre-K programming through elementary school and middle school (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2014;

Muschkin et al., 2015), findings from the current study point out the conditional nature of these effects. We found that the quality of educational opportunity during school made a difference in determining whether ECE program effects on student achievement were maintained or diminished in the long run. Moreover, we found that the conditions were varied—including evidence of dynamic complementarity between Smart Start and subsequent educational opportunity as well as evidence of dynamic substitutability between NC Pre-K and subsequent educational opportunity. Today, educators and policymakers recognize the importance of promoting continuity in high-quality education across the lifespan, and several frameworks have been developed to facilitate educational alignment (e.g., Kauerz & Coffman, 2013; Stipek et al., 2017; Zellman & Kilburn, 2015). Consistent with these alignment frameworks, we advocate for a view of ECE programming and its benefits as part of a continuum of high-quality educational inputs to be built upon sequentially across development (Carr, 2021).

Our study was unique in focusing on the quality of educational opportunity in school districts as the unit of analysis. We found that school-district average achievement and achievement growth both had positive main effects on student reading and math achievement in eighth grade. However, it is important to note that average achievement—but not achievement growth—showed a reliable, negative association with a measure of school-district socioeconomic disadvantage. This suggests that average achievement may index a wide range of educational opportunities available to students in their communities. Alternatively, achievement growth may better reflect the unique contributions of schooling toward promoting student skill development. Therefore, these different dimensions of educational opportunity may represent different pathways toward raising student achievement. For example, the benefits of enrolling in a school district with higher levels of average achievement may begin to accrue as soon as

children step inside the kindergarten door—impacting early skill development—while the benefits of achievement growth may require multiple years of exposure to accrue benefits—impacting later skill development. Additionally, the educational benefits of enrolling in a school district with higher levels of average achievement may extend beyond the experiences students have during the school day to their experiences in community centers and after school programs.

Finally, although the effects found in this study may be considered small by conventional standards (e.g., a 0.017 *SD* increase in eighth grade reading scores with every 1 *SD* unit increase in NC Pre-K funding exposure), these effects were educationally meaningful according to empirical benchmarks established for educational interventions (Kraft, 2020). These small magnitude effects can be further contextualized by the moderate cost (i.e., not small or large; Kraft, 2020) of the average level of Smart Start and NC Pre-K program exposure ($1,539 and $1,064, respectively). Overall, reasonable cost investments made to the Smart Start and NC Pre-K programs produced population-wide improvements in student achievement outcomes that remained educationally meaningful in the long-term—suggesting that even larger investments could produce greater impact in the future.

Study Limitations

The current study focused on school districts as the unit of analyses to measure the quality of educational opportunity, and our measures of average achievement and achievement growth were both positively associated with student achievement in eighth grade. Nonetheless, it is possible that the quality of educational opportunity may vary between schools within school districts. However, this consideration was outside the scope of our study because many schools did not include the full range of grade levels needed to measure average achievement and achievement growth from third to eighth grade. Additionally, we did not examine moderation of

ECE program exposure effects in earlier grades, because school-district average achievement and achievement growth entailed measurement of student achievement between third and eighth grade—confounding earlier grade outcomes with information from eighth grade academic achievement.

Conclusion

Our study adds to a growing body of evidence documenting the population-wide improvements in student achievement resulting from financial investments in ECE programs (e.g., Fitzpatrick, 2008; Kose, 2021). We confirmed prior research to indicate that financial investments NC's two, flagship ECE programs—Smart Start and NC Pre-K—can have population-wide effects on student academic achievement that endure through the end of middle school—nine years after program exposure (Bai et al., 2020). The current study also contributes new evidence regarding the role of school-age educational opportunity in differentiating these long-term effects. Specifically, we found that the positive effect of financial investments in Smart Start was larger for students who subsequently enrolled in higher-quality school districts, while the positive effect of NC Pre-K was smaller for students who subsequently enrolled in higher-quality school districts. The existing literature in this area was quite mixed to begin with (see Bailey, Jenkins, et al., 2020), and our findings provide support for two alternative hypotheses—including dynamic complementarity and dynamic substitutability—as mechanisms by which the quality of subsequent educational opportunity during school may moderate the long-term benefits of ECE programming. In the debate about fadeout or persistence of ECE program effects on student achievement, we encourage educational stakeholders and policymakers to consider how subsequent schooling environments can play a role in determining whether program effects are maintained or diminished in the long run.

References

Abenavoli, R. M. (2019). The mechanisms and moderators of "fade-out": Towards understanding why the skills of early childhood program participants converge over time with the skills of other children. *Psychological Bulletin, 145*(12), 1103–1127. https://doi.org/10.1037/bul0000212

Andrabi, T., Das, J., Ijaz Khwaja, A., & Zajonc, T. (2011). Do value-added estimates add value? Accounting for learning dynamics. *American Economic Journal: Applied Economics, 3*(3), 29–54. https://doi.org/10.1257/app.3.3.29

Ansari, A., & Pianta, R. C. (2018). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review, 86*(1), 120–127. https://doi.org/10.1016/j.childyouth.2018.01.025

Bai, Y., Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2020). Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow? *Children and Youth Services Review, 112*(1), 1–10. https://doi.org/10.1016/j.childyouth.2020.104890

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest, 21*(2), 55–97. https://doi.org/10.1177/1529100620915848

Bailey, D. H., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness, 10*(1), 7–39. https://doi.org/10.1080/19345747.2016.1232459

Bailey, D. H., Jenkins, J. M., & Alvarez-Vargas, D. (2020). Complementarities between early educational intervention and later educational quality? A systematic review of the

sustaining environments hypothesis. *Developmental Review, 56*(1), 1–22. https://doi.org/10.1016/j.dr.2020.100910

Bailey, M. J., Sun, S., & Timpe, B. D. (2021). Prep school for poor kids: The long-run impacts of Head Start on human capital and economic self-sufficiency. *NBER Working Paper No. 28268*. http://www.nber.org/papers/w28268

Barnett, W. S., Epstein, D. J., Carolan, M. E., Fitzgerald, J., Ackerman, D. J., & Friedman, A. H. (2010). *The State of Preschool 2010: State Preschool Yearbook*. National Institute for Early Education Research. https://nieer.org

Barnett, W. S., & Jung, K. (2021). Effects of New Jersey's Abbott preschool program on children's achievement, grade retention, and special education through tenth grade. *Early Childhood Research Quarterly, 56*(1), 248–259. https://doi.org/10.1016/j.ecresq.2021.04.001

Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., Howes, C., & Daniel-Echols, M. (2018). State prekindergarten effects on early learning at kindergarten entry: An analysis of eight state programs. *AERA Open, 4*(2), 1–16. https://doi.org/10.1177/2332858418766291

Barnett, W. S., Robin, K. B., Hustedt, J. T., & Schulman, K. L. (2003). *The state of preschool 2003: State preschool yearbook.* National Institute for Early Education Research. https://nieer.org

Bassok, D., Gibbs, C. R., & Latham, S. (2019). Preschool and children’s outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? *Child Development, 90*(6), 1–23. https://doi.org/10.1111/cdev.13067

Bazemore, M., & Van Dyk, P. B. (2004). *North Carolina reading comprehension tests: Technical report (Ed.2)*. North Carolina Department of Public Instruction.

Bierman, K. L., Nix, R. L., Heinrichs, B. S., Domitrovich, C. E., Gest, S. D., Welsh, J. A., & Gill, S. (2014). Effects of Head Start REDI on children's outcomes 1 year later in different kindergarten contexts. *Child Development, 85*(1), 140–159. https://doi.org/10.1111/cdev.12117

Bryant, D., Maxwell, K., Taylor, K., Poe, M., Peisner-Feinberg, E., & Bernier, K. (2003). *Smart Start and preschool child care quality in North Carolina: Change over time and relation to children's readiness*. The University of North Carolina.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, Inc.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. The University of Chicago Press.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record, 112*(3), 579–620. https://doi.org/10.1177/016146811011200303

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002, 2002/01/01). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science, 6*(1), 42–57. https://doi.org/10.1207/S1532480XADS0601_05

Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy, 6*(4), 135–173. https://doi.org/10.1257/pol.6.4.135

Carr, R. C. (2021). *The benefits of early childhood education can persist in the long run*. The Hunt Institute.

Carr, R. C., Peisner-Feinberg, E. S., Kaplan, R., & Mokrova, I. L. (2021). Effects of North Carolina's pre-kindergarten program at the end of kindergarten: Contributions of school-wide quality. *Applied Developmental Psychology, 76*(1). https://doi.org/10.1016/j.appdev.2021.101317

Chenoweth, K. (2017). *Schools that succeed: How educators marshal the power of systems for improvement*. Harvard Education Press.

Chenoweth, K. (2021). *Districts that succeed: Breaking the correlation between race, poverty, and achievement*. Harvard Education Press.

Cohen-Vogel, L., Sadler, J., Little, M. H., Merrill, B., & Curran, F. C. (2020). The adoption of public pre-kindergarten among the American states: An event history analysis. *Educational Policy, 0*(0), 1–33. https://doi.org/10.1177/0895904820961002

Cunha, F., Heckman, J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence of life cycle skill formation. In F. Hanushek & F. Welch (Eds.), *Handbook of the economics of education*. Elsevier.

Curenton, S. M., Dong, N., & Shen, X. (2015, Jul). Does aggregate school-wide achievement mediate fifth grade outcomes for former early childhood education participants? *Developmental Psychology, 51*(7), 921–934. https://doi.org/10.1037/a0039295

Currie, J., & Thomas, D. (1995). Does Head Start Make a Difference? *The American Economic Review, 85*(3), 341–364. http://www.jstor.org/stable/2118178

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2016, Nov 17). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary school. *Child Development, 88*(3), 996–1014. https://doi.org/10.1111/cdev.12645

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D.-M. C., Crawford, G. M., Frome, P. M., Clifford, R. M., Burchinal, M., Howes, C., Bryant, D., & Pianta, R. C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 25*(2), 177–193. https://doi.org/10.1016/j.ecresq.2009.10.003

Eccles, J. S., & Roeser, R. W. (2010). An ecological view of schools and development. In J. S. Eccles & J. L. Meece (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 6–21). Routledge.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Policy, 8*(1), 1–38. https://doi.org/doi:10.2202/1935-1682.1897

Friedman-Krauss, A. H., Barnett, S. W., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., & Gardiner, B. A. (2020). *The state of preschool 2019: State preschool yearbook*. National Institute for Early Education Research. https://nieer.org

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *The American Economic Review, 92*(4), 999–1012. https://doi.org/https://www.jstor.org/stable/3083291

Gormley, W. T., Phillips, D., & Anderson, S. (2018). The Effects of Tulsa's Pre-K Program on Middle School Student Performance. *Journal of Policy Analysis and Management, 37*(1), 63-87. https://doi.org/10.1002/pam.22023

Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320*(5884), 1723–1724. https://doi.org/10.1126/science.1156019

Gray-Lobe, G., Pathak, P., & Walters, C. (2021). The long-term effects of universal preschool in Boston. *SEII Discussion Paper #2021.05*.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, *312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Hill, C. J., Gormley, W. T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly, 32*(3), 60–79. https://doi.org/10.1016/j.ecresq.2014.12.005

Johnson, R. C., & Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending. *American Economic Journal: Economic Policy, 11*(4), 310–349. https://doi.org/10.1257/pol.20180510

Kauerz, K., & Coffman, J. (2013). *Framework for planning, implementing, and evaluating prek-3rd grade approaches*. National P-3 Center.

Kose, E. (2021). Public investments in early childhood education and academic performance: Evidence from Head Start in Texas. *Journal of Human Resources, 0*(00). https://doi.org/10.3368/jhr.0419-10147R2

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. https://doi.org/10.3102/0013189x20912798

Kropko, J., & Kubinec, R. (2020). Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PLoS ONE*, *15*(4), 1–22. https://doi.org/10.1371/journal.pone.0231349

Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. *Journal of Policy Analysis and Management*, *33*(1), 162–187. https://doi.org/10.1002/pam.21734

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*(1), 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005

Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, *122*(1), 159–208. http://www.jstor.org/stable/25098840

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly*, *22*(1), 18–38. https://doi.org/10.1016/j.ecresq.2006.10.002

Mbella, K., Zhu, M., Karkee, T., & Lung, H. (2016a). *The North Carolina testing program technical report, 2012–2015: English language arts/reading assessments (ELA) end-of-grade 3–8 and end-of-course English II*. Department of Public Instruction.

Mbella, K., Zhu, M., Karkee, T., & Lung, H. (2016b). *The North Carolina testing program technical report, 2012–2015: Mathematics assessments end-of-grade 3–8 and end-of-course Math I*. Department of Public Instruction.

McCartney, K., Dearing, E., Taylor, B. A., & Bub, K. L. (2007). Quality child care supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology*, *28*(5-6), 411–426. https://doi.org/10.1016/j.appdev.2007.06.010

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., Yang, R., Koepp, A., & Shonkoff, J. P. (2017). Impacts of early childhood education on medium- and long-term educational outcomes. *Educational Researcher, 46*(8), 474–487. https://doi.org/10.3102/0013189x17737739

Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's early childhood initiatives on special education placements in third grade. *Educational Evaluation and Policy Analysis, 37*(4), 478–500. https://doi.org/10.3102/0162373714559096

North Carolina Department of Public Instruction. (2009). *North Carolina reading comprehension tests: Technical report (Ed. 3)*. Author.

North Carolina Partnership for Children. (2010). *Smart Start impact report 2009-2010*. https://www.smartstart.org

North Carolina Partnership for Children. (2019). *Annual report to the North Carolina General Assembly for fiscal year 2018-2019*. https://www.smartstart.org

Pages, R., Lukes, D. J., Bailey, D. H., & Duncan, G. J. (2020). Elusive longer-run impacts of Head Start: Replications within and across cohorts. *Educational Evaluation and Policy Analysis, 42*(4), 471–492. https://doi.org/10.3102/0162373720948884

Pearman, F. A., Springer, M., Lipsey, M., Lachowicz, M., Swain, W., & Farran, D. (2020). Teachers, schools, and pre-K effect persistence: An examination of the sustaining environment hypothesis. *Journal of Research on Educational Effectiveness*, *13*(4), 1–27. https://doi.org/10.1080/19345747.2020.1749740

Peisner-Feinberg, E., Kuhn, L., Zadrozny, S., Foster, T., & Burchinal, M. (2020). *Kindergarten follow-up findings from a small-scale RCT study of the North Carolina Pre-Kindergarten Program*. The University of North Carolina at Chapel Hill.

Peisner-Feinberg, E. S., Mokrova, I. L., & Anderson, T. L. (2017). *Effects of participation in the North Carolina Pre-Kindergarten Program at the end of kindergarten 2015-2016 statewide evaluation*. In. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., Schaaf, J., Hildebrandt, L., & LaForett, D. (2013). *Quality and characteristics of the North Carolina Pre-Kindergarten program: 2011-2012 statewide evaluation*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2010). *Long-term effects of the North Carolina More at Four Pre-kindergarten Program: Children's reading and math skills at third grade: Full report*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Effects of the North Carolina More at Four Pre-kindergarten Program on children's school readiness skills: Summary of key findings*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L. M., & Sideris, J. (2014). *Effects of Georgia's Pre-K Program on children's school readiness skills: Findings from the 2012–2013 evaluation study.* The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., Zadrozny, S., Kuhn, L., & Van Manen, K. (2019). *Effects of the North Carolina Pre-Kindergarten Program: Findings through Pre-K of a Small-Scale RCT Study*. The University of North Carolina, FPG Child Development Institute.

Phillips, D., Gormley, W., & Anderson, S. (2016). The effects of Tulsa's CAP Head Start program on middle-school academic outcomes and progress. *Developmental Psychology, 52*(8), 1247–1261. https://doi.org/10.1037/dev0000151

Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J., Dynarski, M., Magnuson, K. A., & Weiland, C. (2017). *The current state of scientific knowledge on pre-kindergarten effects.* Brookings Institution.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study: Final report*. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., Ciarico, J., Rohacek, M., Adams, G., & Spier, E. (2010). *Head Start Impact Study: Final report*. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Reardon, S. F. (2018). *Educational opportunity in early and middle childhood: Variation by place and age* (CEPA Working Paper No.17-12). Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp17-12

Rossin-Slater, M., & Wüst, M. (2020). What is the added value of preschool for poor children? Long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics, 12*(3), 255–286. https://doi.org/10.1257/app.20180698

Sanford, E. E. (1996). *North Carolina end-of-grade tests technical report # 1: Reading comprehension, mathematics*. Department of Public Instruction.

Schweinhart, L. J., Berrueta-Clement, J. R., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1985). Effects of the Perry Preschool Program on youths through age 19: A summary. *Topics in Early Childhood Special Education, 5*(2), 26–35. https://doi.org/10.1177/027112148500500204

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Stipek, D., Clements, D., Coburn, C., Franke, M., & Farran, D. (2017). PK-3: What does it mean for instruction? *SRCD Social Policy Report, 30*(2), 1–23.

Swain, W. A., Springer, M. G., & Hofer, K. G. (2015). Early grade teacher effectiveness and pre-k effect persistence: Evidence from Tennessee. *AERA Open, 1*(4), 1–17. https://doi.org/10.1177/2332858415612751

Thompson, O. (2018). Head Start's long-run impact. *Journal of Human Resources, 53*(4), 1100–1139. https://doi.org/10.3368/jhr.53.4.0216.7735R1

Unterman, R., & Weiland, C. (2020). Higher-quality elementary schools sustain the prekindergarten boost: Evidence from an exploration of variation in the Boston prekindergarten program's impacts. (EdWorkingPaper: 20-321). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/hkv0-wj29

Vincent, P., R. (2018). Time trends matter: The case of medical cannabis laws and opioid overdose mortality. *MPRA Paper No. 87237*.

Weiland, C., Unterman, R., Shapiro, A., Staszak, S., Rochester, S., & Martin, E. (2019). The effects of enrolling in oversubscribed prekindergarten programs through third grade. *Child Development, 91*(5), 1401–1422. https://doi.org/10.1111/cdev.13308

Weiland, C., & Yoshikawa, H. (2013, Nov-Dec). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills [Research Support, Non-U.S. Gov't]. *Child Development*, *84*(6), 2112-2130. https://doi.org/10.1111/cdev.12099

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gromley, W. T., Ludwig, J., Magnuson, K. A., Phillips, D., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Foundation for Child Development.

Zellman, G., & Kilburn, M. (2015). *How well did P–3 work? The Hawai'i preschool-third grade education reform initiative.* Rand Corporation.

Zhai, F., Raver, C. C., & Jones, S. M. (2012, May 01). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. *Children and Youth Services Review*, *34*(5), 946–954. https://doi.org/10.1016/j.childyouth.2012.01.026

Table 1

Descriptive Statistics for the Eighth Grade Student Reading Achievement Analysis Sample (N = 1,164,948)

Variable	*M*	(*SD*)
Program funding		
Smart Start funding exposure (in $00's; all study years)	14.56	(10.90)
NC Pre-K funding exposure (in $00's; all study years)	4.18	(6.71)
Eighth grade achievement outcomes		
Reading	-0.01	(1.00)
Math	0.00	(1.00)
School-district average achievement		
Reading	-0.09	(0.18)
Math	-0.10	(0.19)
School-district achievement growth		
Reading	-0.02	(0.10)
Math	0.00	(0.16)
Student characteristics		
Sex (Male)	0.50	(0.50)
Extremely low birth weight	0.00	(0.06)
Very low birth weight	0.01	(0.09)
Low birth weight	0.07	(0.25)
Normal birth weight	0.82	(0.38)
High birth weight	0.10	(0.29)
Non-Hispanic White	0.58	(0.49)
African American	0.29	(0.45)
Native American	0.02	(0.13)
Asian	0.01	(0.11)
Hispanic	0.07	(0.25)
Mixed race	0.03	(0.17)
Maternal education (years)	12.58	(2.55)
Parent marital status	0.65	(0.48)
Maternal age (years)	26.11	(5.93)
No dad information	0.14	(0.34)
Mother immigrant	0.09	(0.29)
First born	0.43	(0.50)

Table 1 (continued)

Descriptive Statistics for the Eighth Grade Student Reading Achievement Analysis Sample (N = 1,164,948)

Variable	*M*	(*SD*)
Mother White	0.68	(0.47)
Mother African American	0.29	(0.45)
Mother Native American	0.02	(0.12)
Mother Asian	0.01	(0.12)
Mother Hispanic	0.07	(0.26)
Mother other race	0.00	(0.03)
School-district characteristics, eighth-grade year		
Total per pupil expenditures (2019 $s)	9821	(976)
Economic disadvantage (%)	52.00	(15.76)
School district membership	39653	(45491)
County characteristics, birth year		
Births to black mothers (% of births)	25.75	(14.46)
Births to Hispanic mothers (% of births)	8.03	(11.37)
Births to low education mothers (% of births)	22.59	(6.56)
Population on Food Stamps (% of population)	7.40	(3.64)
Population on Medicaid (% of population)	14.40	(5.90)
Number of births (log)	7.51	(1.09)
Total population (log)	11.77	(1.00)
Median family income (2019 $s)	68289	(13223)

Note. Descriptive statistics are provided for those students included in the reading outcome analyses (N = 1,164,948). Descriptive statistics for students included in the math outcome analyses are provided in Table S4. The Smart Start and NC Pre-K funding variables were scaled such that a value of 1.00 corresponds to $100 in funding. School-district average achievement in reading/math (β_{00j}) was measured at third grade and achievement growth in reading/math (β_{10j}) was measured from third to eighth grade. All values of school-district average achievement and achievement growth were "current year minus 1." Economic disadvantage was indexed based on the percent of students in the school district who qualified for free- or reduced-price lunch

Table 2

Eighth-Grade Student Achievement Analysis Results

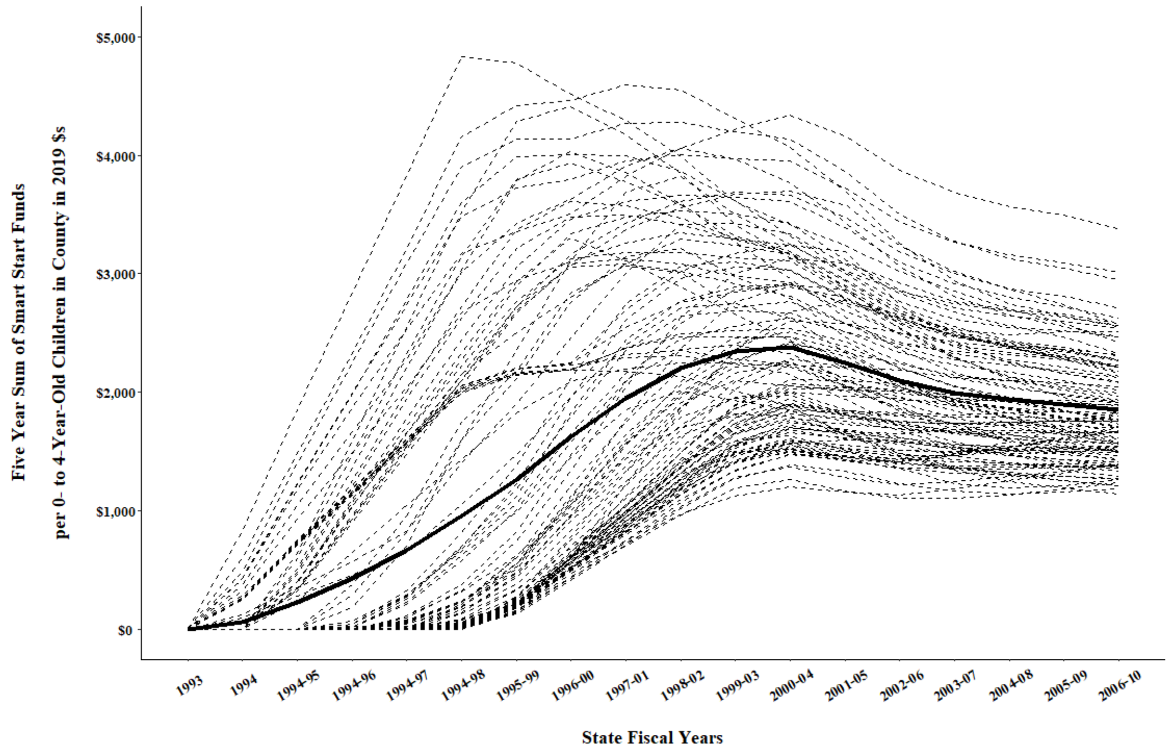
	Reading				Math			
	N = 1,164,948				*N* = 1,127,587			
Parameter	*β*	(*SE*)	*β*	(*SE*)	*β*	(*SE*)	*β*	(*SE*)
Smart Start funding exposure	0.012†	(0.007)	0.015*	(0.006)	0.009	(0.008)	0.035**	(0.011)
NC Pre-K funding exposure	0.017*	(0.008)	0.021**	(0.008)	0.020†	(0.012)	–0.003	(0.010)
Average Achievement	0.062***	(0.005)	0.061***	(0.005)	0.083***	(0.005)	0.081***	(0.005)
Smart Start × Average Achievement	0.007**	(0.002)	0.009***	(0.002)	0.003	(0.003)	0.006*	(0.002)
NC Pre-K × Average Achievement	–0.003	(0.002)	–0.004†	(0.002)	0.001	(0.003)	–0.005*	(0.002)
Achievement Growth	0.037***	(0.003)	0.036***	(0.003)	0.093***	(0.004)	0.089***	(0.004)
Smart Start × Achievement Growth	0.003	(0.002)	0.002	(0.002)	0.006†	(0.003)	0.007*	(0.003)
NC Pre-K × Achievement Growth	–0.003*	(0.001)	–0.005**	(0.001)	0.003	(0.003)	–0.001	(0.003)
Student, school-district, & county covariates	Inc.		Inc.		Inc.		Inc.	
County fixed effects	Inc.		Inc.		Inc.		Inc.	
Program year fixed effects	Inc.		Inc.		Inc.		Inc.	
School year fixed effects	Inc.		Inc.		Inc.		Inc.	
County/school year linear time trends	Inc.		Inc.		Inc.		Inc.	
County/school year quadratic time trends			Inc.				Inc.	
R-squared	0.29		0.30		0.29		0.30	

Note. Parameter estimates for the student, school-district, and county covariates; county fixed effects; program year fixed effects; school year fixed effects; and county/school year linear and quadratic time trends are not displayed in Table 2, but can be made available from the corresponding author upon request. School-district average achievement/achievement growth correspond to reading average achievement/achievement growth in the reading models and math average achievement/achievement growth in the math models. All values of school-district average achievement and achievement growth were “current year minus 1.” The student outcome

variables, the Smart Start and NC Pre-K funding exposure variables, and the school-district average achievement and achievement growth variables were standardized with $M = 0$, $SD = 1$. All models were estimated with robust standard errors clustered at the county level. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Figure 1

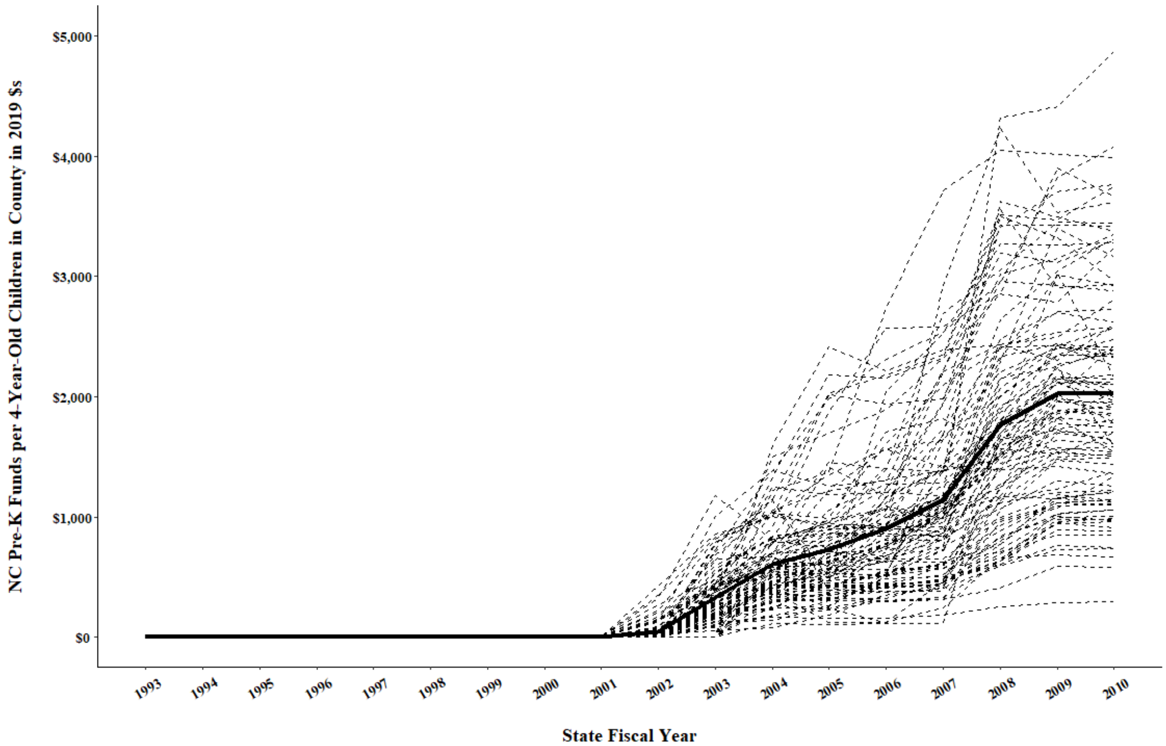
Five Year Sum of Smart Start Funds per 0- to 4-Year-Old Children in County: 1993 to 2010



Note. All counties with statewide average. Dotted lines represent individual counties. Solid line represents statewide average. Smart Start was established in state fiscal year 1994. In the student outcome analyses, all counties were assigned a pre-trend value of "0" prior to the establishment of Smart Start during state fiscal year 1994.

Figure 2

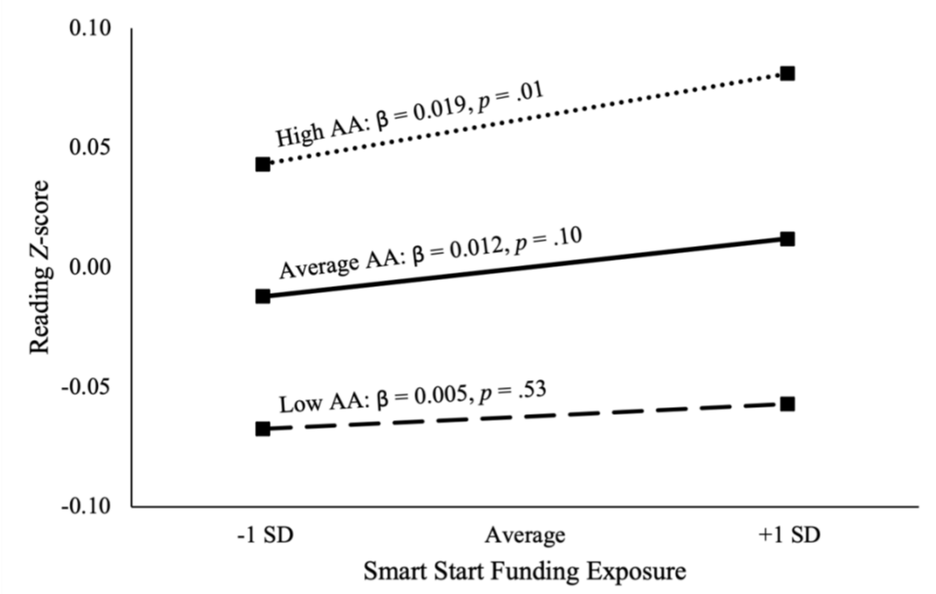
NC Pre-K Funds per 4-Year-Old Children in County: 1993 to 2010



Note. All counties with statewide average. Dotted lines represent individual counties. Solid line represents statewide average. NC Pre-K was established in state fiscal year 2002. In the student outcome analyses, all counties were assigned a pre-trend value of "0" prior to the establishment of NC Pre-K in state fiscal year 2002.

Figure 3

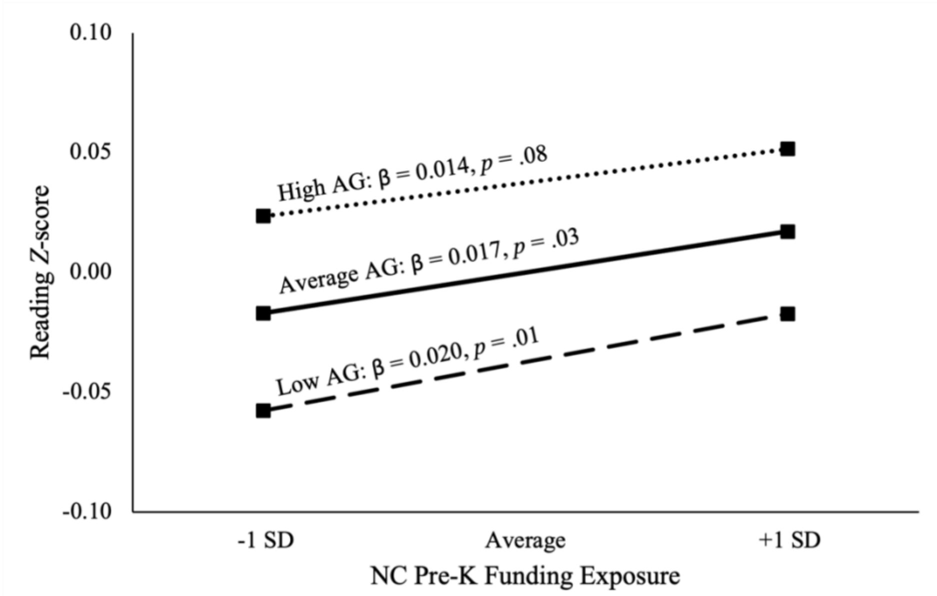
Moderation of the Smart Start Funding Effect on Eighth Grade Student Reading Achievement by School-District Average Achievement in Reading



Note. AA = Average achievement. The effect of Smart Start funding exposure on student reading outcomes in eighth grade at three values of school-district average achievement in reading: 1 *SD* below the average (low), at the average, and 1 *SD* above the average (high).

Figure 4

Moderation of the NC Pre-K Funding Effect on Eighth Grade Student Reading Achievement by School-District Achievement Growth in Reading



Note. AG = Achievement growth. The effect of NC Pre-K funding exposure on student reading outcomes in eighth grade at three values of school-district achievement growth in reading: 1 *SD* below the average (low), at the average, and 1 *SD* above the average (high).

Supplemental Materials

Table S1

Analyses to Examine the Association between Program Funding Exposure and School-District Average Achievement & Achievement Growth (N = 2,070)

	Reading				Math			
	Average Achievement		Achievement Growth		Average Achievement		Achievement Growth	
Parameter	β	(*SD*)	β	(*SD*)	β	(*SD*)	β	(*SD*)
Smart Start funding exposure	0.05	(0.04)	–0.12	(0.08)	0.03	(0.05)	–0.11	(0.08)
NC Pre-K funding exposure	0.01	(0.03)	0.11	(0.08)	–0.08	(0.05)	0.14	(0.09)

Note. The generalized equation for these analyses was:

$$O_{scPy} = \beta_0 + \beta_1 Smart\ Start\ funding_{cPy} + \beta_2 NC\ PreK\ funding_{cPy} + \alpha_c + \gamma_{Py} + \epsilon_{scPy}$$

O is the school-district average achievement or achievement growth score for school district *s* in county *c* aligned to program year *Py*, β_1 is the main effect of Smart Start funding, β_2 is the main effect of NC Pre-K funding, α is the fixed effect for county, and γ is the fixed effect for program year. County-level program funding information in program year *T* was aligned with school-district average achievement and achievement growth information during school year *T* + 4 (e.g., students' exposure to program funding in pre-K during program year 2000 was aligned to the average achievement and achievement growth values for students in third grade four years later during school year 2004, which is the school year in which these students were expected to be in third grade). Additionally, all values of school-district average achievement and achievement growth were "current year minus 1." The Smart Start and NC Pre-K funding variables as well as the school-district average achievement and achievement growth variables were standardized with $M = 0$, $SD = 1$. All models were estimated with robust standard errors clustered at the county level. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table S2

Descriptive Statistics for School-District Average Achievement and Achievement Growth Presented at the School-District Level (N = 2,070)

Variable	*M*	(*SD*)
School-district average achievement		
Reading	-0.13	(0.22)
Math	-0.14	(0.23)
School-district achievement growth		
Reading	-0.02	(0.12)
Math	-0.01	(0.20)

Note. School-district average achievement in reading/math (β_{00j}) was measured at third grade and achievement growth in reading/math (β_{10j}) was measured from third to eighth grade. Economic disadvantage was indexed based on the percent of students in the school district who qualified for free- or reduced-price lunch. There were 115 school districts in each of the 18 cohorts (i.e., 18 × 15 = 2,070 school district by cohort observations).

Table S3

Analyses to Examine the Association between School-District Economic Disadvantage and School-District Average Achievement & Achievement Growth (N = 2,070)

	Reading				Math			
	Average Achievement		Achievement Growth		Average Achievement		Achievement Growth	
Parameter	β	(*SD*)	β	(*SD*)	β	(*SD*)	β	(*SD*)
Economic disadvantage (%)	–0.46***	(0.08)	–0.09	(0.07)	–0.39***	(0.10)	–0.09	(0.08)

Note. Economic disadvantage is the percent of economically disadvantaged students in the school district (i.e., the percent of students in the school district who qualified for free- or reduced-price lunch). The generalized equation for these analyses was:

$$O_{scSy} = \beta_0 + \beta_1 Economic\ Disadvantage_{scSy} + \alpha_c + \gamma_{Sy} + \epsilon_{scSy}$$

O is the school-district average achievement or achievement growth score for school district *s* in county *c* aligned to school year *Sy*, β_1 is the main effect of economic disadvantage, α is the fixed effect for county, and γ is the fixed effect for school year. School-district economic disadvantage information in school year *Sy* was aligned with school-district average achievement and achievement growth information for the cohort of students in third grade during school year *Sy* (e.g., school-district economic disadvantage in school year 2000 was aligned to the average achievement and achievement growth values for the cohort of students in third grade in 2000). The school-district economic disadvantage, average achievement, and achievement growth variables were standardized with $M = 0$, $SD = 1$. All models were estimated with robust standard errors clustered at the county level. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table S4

Descriptive Statistics for the Eighth Grade Student Math Achievement Analysis Sample (N = 1,127,587)

Variable	*M*	(*SD*)
Program funding		
Smart Start funding exposure (in $00's; all study years)	14.46	(11.05)
NC Pre-K funding exposure (in $00's; all study years)	3.84	(6.39)
Eighth grade achievement outcomes		
Reading	-0.04	(1.00)
Math	0.00	(1.00)
School-district average achievement		
Reading	-0.09	(0.18)
Math	-0.10	(0.19)
School-district achievement growth		
Reading	-0.02	(0.10)
Math	-0.01	(0.16)
Student characteristics		
Sex (Male)	0.50	(0.50)
Extremely low birth weight	0.00	(0.06)
Very low birth weight	0.01	(0.09)
Low birth weight	0.07	(0.25)
Normal birth weight	0.82	(0.38)
High birth weight	0.10	(0.29)
Non-Hispanic White	0.58	(0.49)
African American	0.30	(0.46)
Native American	0.02	(0.13)
Asian	0.01	(0.10)
Hispanic	0.07	(0.25)
Mixed race	0.03	(0.17)
Maternal education (years)	12.53	(2.53)
Parent marital status	0.65	(0.48)
Maternal age (years)	26.01	(5.91)
No dad information	0.14	(0.35)
Mother immigrant	0.09	(0.28)
First born	0.43	(0.50)

Table S4 (continued)

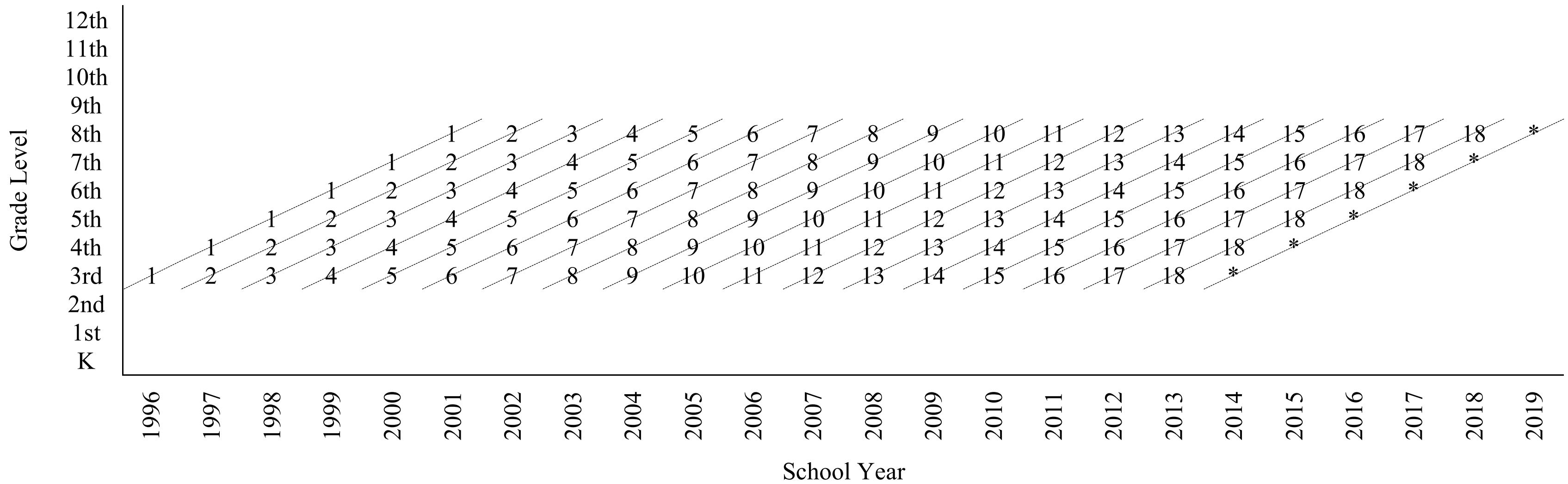
Descriptive Statistics for the Eighth Grade Student Math Achievement Analysis Sample (N = 1,127,587)

Variable	*M*	(*SD*)
Mother White	0.67	(0.47)
Mother African American	0.30	(0.46)
Mother Native American	0.02	(0.13)
Mother Asian	0.01	(0.11)
Mother Hispanic	0.07	(0.26)
Mother other race	0.00	(0.03)
School-district characteristics, eighth-grade year		
Total per pupil expenditures (2019 $s)	9817	(978)
Economic disadvantage (%)	51.88	(15.64)
School district membership	38842	(44607)
County characteristics, birth year		
Births to black mothers (% of births)	25.85	(14.53)
Births to Hispanic mothers (% of births)	7.77	(11.43)
Births to low education mothers (% of births)	22.60	(6.58)
Population on Food Stamps (% of population)	7.38	(3.65)
Population on Medicaid (% of population)	14.32	(5.90)
Number of births (log)	7.49	(1.08)
Total population (log)	11.75	(1.00)
Median family income (2019 $s)	67998	(13132)

Note. Descriptive statistics are provided for those students included in the math outcome analyses (N = 1,127,587). The Smart Start and NC Pre-K funding variables were scaled such that a value of 1.00 corresponds to $100 in funding. School-district average achievement in reading/math (β_{00j}) was measured at third grade and achievement growth in reading/math (β_{10j}) was measured from third to eighth grade. All values of school-district average achievement and achievement growth were "current year minus 1." Economic disadvantage was indexed based on the percent of students in the school district who qualified for free- or reduced-price lunch

Figure S1

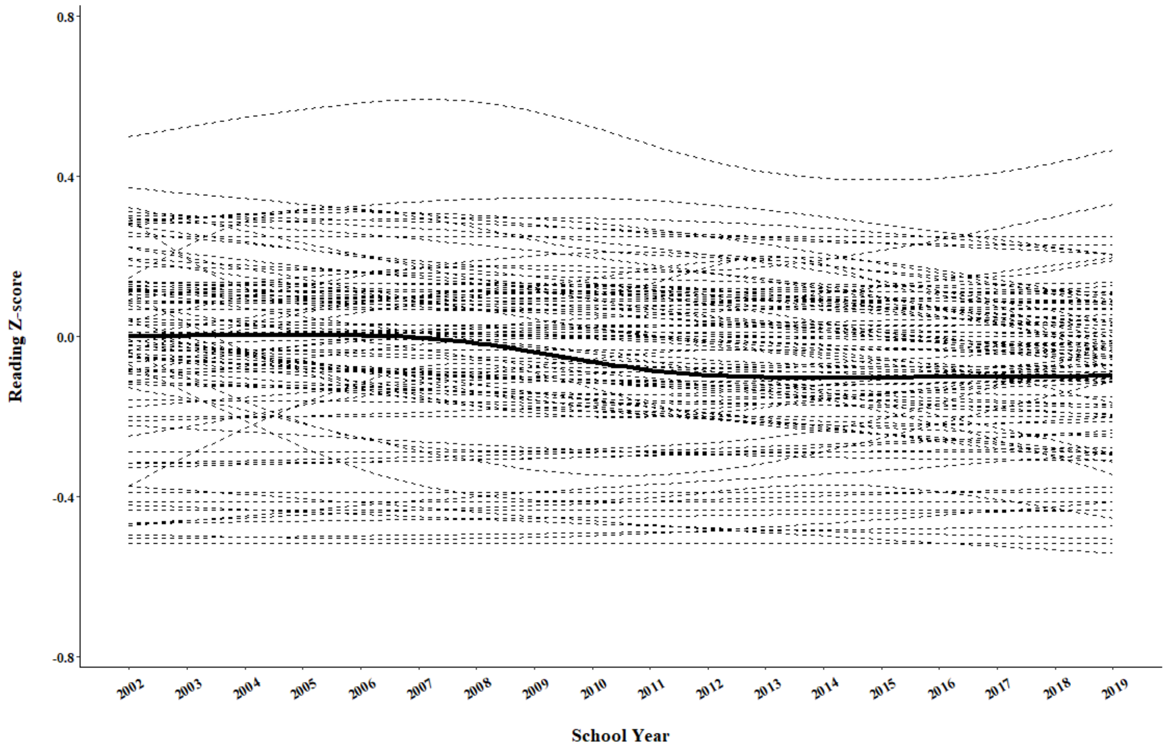
North Carolina Third to Eighth Grade Student Cohorts between 1996 and 2018 (N = 18)



Note. Each diagonal line corresponds to a cohort of students. In Phase I of our analyses, a separate growth curve model was estimated for students in each cohort, which was used to calculate the intercept (i.e., third grade average achievement) and slope (i.e., achievement growth between third and eighth grade) in reading and mathematics for each school district within each cohort. These estimates were then used in Phase II of our analyses to examine moderation of the long-term effects of Smart Start and NC Pre-K. Students in the cohort indexed by the asterisk were included in the Phase II analyses, but not the Phase I analyses because of the "current year minus 1" approach.

Figure S2

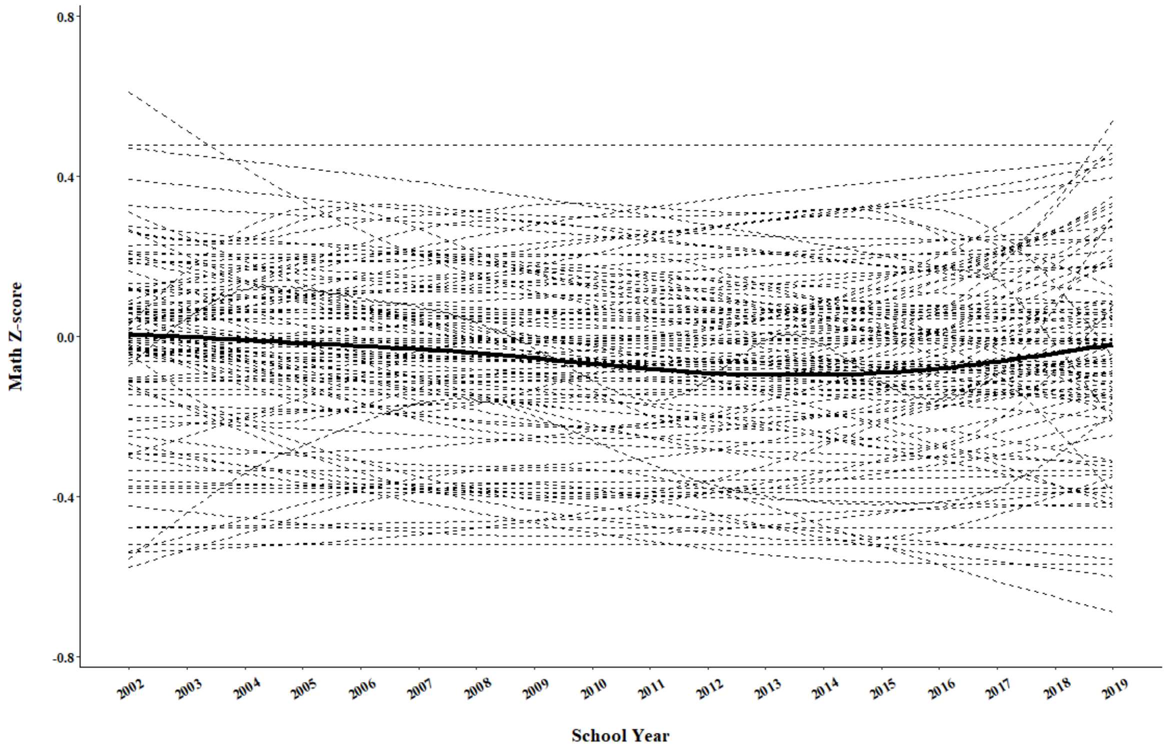
County-Level Means of Eighth-Grade Reading Z-scores: School Years 2002–2019



Note. All 100 counties with statewide average. Dotted lines represent the individual counties. Solid line represents statewide average. A smoothing function (the generalized additive model) was applied to the data.

Figure S3

County-Level Means of Eighth-Grade Math Z-scores: School Years 2002–2019



Note. All 100 counties with statewide average. Dotted lines represent the individual counties. Solid line represents statewide average. A smoothing function (the generalized additive model) was applied to the data.